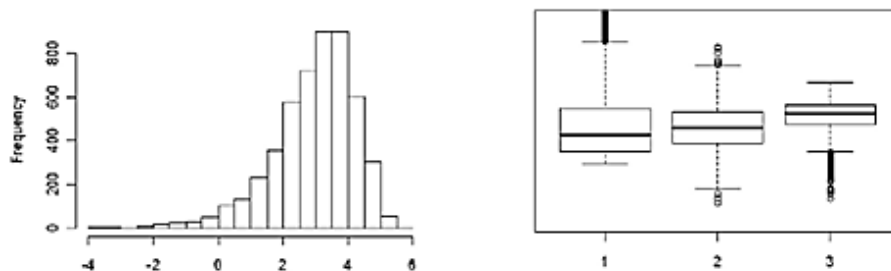


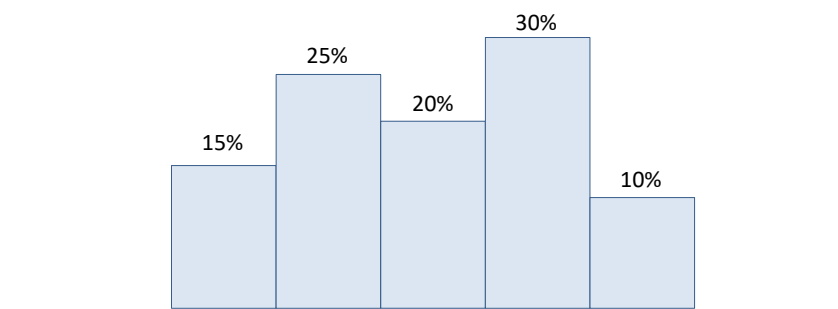
ET585 – Análise Exploratória de Dados

Prof. Dr. Vinícius Quintas Souto Maior
Lista de Exercícios – Análise Exploratória de Dados

Exercício 1 - Na figura abaixo apresenta o histograma para um conjunto de dados. Qual dos *box-plots* é consistente com o histograma na descrição dos dados. Justifique a sua resposta.



Exercício 2 – Dado o histograma abaixo, calcular a média, a variância, a moda, a mediana e o 1º quartil.



Exercício 3 – Em uma granja foi observada a distribuição dos frangos em relação ao peso, que era a seguinte:

Peso (em gramas)	n_i
960 - 980	60
980 - 1.000	160
1.000 - 1.020	280
1.020 - 1.040	260
1.040 - 1.060	160
1.060 - 1.080	80

- Qual a média da distribuição?
- Qual a variância da distribuição?
- Construa o histograma.
- Queremos dividir os frangos em quatro categorias, em relação ao peso, de modo que:
 - os 20% mais leves sejam da categoria D;
 - os 30% seguintes sejam da categoria C;
 - os 30% seguintes sejam da categoria B;
 - os 20% seguintes (ou seja, os 20% mais pesados) sejam da categoria A.Quais os limites de peso entre as categorias A, B, C e D?
- O granjeiro decide separar deste lote os animais com peso inferior a dois desvios padrões abaixo da média para receberem ração reforçada, e também separar os animais com peso superior a 1,5 desvios padrões acima da média para usá-los como reprodutores. Qual a porcentagem de animais que serão separados em cada caso?

ET585 – Análise Exploratória de Dados

Exercício 4 – Para distribuição de frequências abaixo construa três histogramas:

Intervalo de classes	Frequência absoluta
0 - 4	28
4 - 9	46
9 - 15	58
15 - 16	20
16 - 17	31
17 - 19	64
19 - 24	149
24 - 60	316
60 - 80	103

- (a) Usando a frequência absoluta.
- (b) Usando a frequência relativa.
- (c) Usando a densidade de frequências.

Exercício 5 – Uma amostra aleatória de 263 compradores de automóveis foi classificada segundo o gênero do comprador, tipo de automóvel adquirido e o estado civil do comprador. Os dados são resumidos a seguir:

	Auto americano	Auto japonês
Homens Casados	41	47
Homens Solteiros	20	36
Mulheres Casadas	42	40
Mulheres Solteiras	12	25

- a) Construa uma tabela de contingência entre gênero do comprador e tipo do automóvel adquirido, faça uma representação gráfica e verifique, por meio de uma medida de associação, se há indícios de diferenças entre as preferências de homens e mulheres.
- b) Construa uma tabela de contingência entre estado civil do comprador e tipo do automóvel adquirido, faça uma representação gráfica e verifique, por meio de uma medida de associação, se há indícios de diferenças entre as preferências de casados e solteiros.
- c) Compare os resultados dos itens (a) e (b). Em qual situação as diferenças de preferência por automóvel americano ou japonês são mais evidentes?

Exercício 6 – Na tabela a seguir, observa-se o tipo de cooperativa e o estado em que está localizada. Verifique se existe associação entre as variáveis. Considere $\alpha = 0,05 = 5\%$.

Estado	Tipo de Cooperativa			
	Consumidor	Produtor	Escola	Outras
São Paulo	514	237	48	119
Paraná	51	102	126	22
Rio Grande do Norte	111	304	139	48

Exercício 7 – Considere os dados abaixo, que são medidas da velocidade do vento no aeroporto da Philadelphia (EUA), às 01:00 h, para os primeiros quinze dias de dezembro de 1974.

22,2	61,1	13,0	27,8	22,2
7,4	7,4	7,4	20,4	20,4
20,1	11,1	13,0	7,4	14,8

Observe o valor atípico 61,1, que na realidade ocorreu devido a forte tempestade do dia 2 de dezembro.

ET585 – Análise Exploratória de Dados

- Faça um gráfico de dispersão unidimensional e marque nele $Q(0,25)$, $Q(0,50)$ e $Q(0,75)$ e comente.
- Construa um ramo-e-folhas e um box-plot. Comente.
- Construa um gráfico de simetria e um gráfico de quantis. Comente sobre a simetria dos dados.

Exercício 8 – Suponha que a faixa etária dos alunos de uma determinada unidade escolar se apresenta segundo a distribuição abaixo. O que você pode dizer com respeito à assimetria e a curtose da distribuição?

Faixa etária (anos)	Número de alunos
7 - 9	197
9 - 11	372
11 - 13	527
13 - 15	114
15 - 17	49
17 - 19	25
19 - 21	3
Total	1.287

Exercício 9 – Considere as seguintes medidas, relativas a três distribuições de frequência:

Distribuição	$Q(0,25)$	$Q(0,75)$	$Q(0,10)$	$Q(0,90)$
A	814	935	772	1012
B	63,7	80,3	55,0	86,6
C	28,8	45,6	20,5	49,8

- Calcule os respectivos graus de curtose.
- Classifique cada uma das distribuições com relação à curva normal.

Exercício 10 – Em uma distribuição de frequência foram encontradas as seguintes medidas: Média = 33,18; Moda = 27,5; Mediana = 31,67; Desvio Padrão = 12,45.

- Classifique o tipo de assimetria.
- Calcule o coeficiente de assimetria.

Exercício 11 – Comente sobre transformações nos dados para obter simetria.